Finite Gelfand pairs and Markov chain Monte-Carlo method (Representation Theory and Combinatorics)

吉田, 知行

数理解析研究所講究録
第 1689 巻 2010 年 164-170

Finite Gelfand pairs and Markov chain Monte-Carlo method

YOSHIDA, Tomoyuki (吉田知行)
yoshidat@math.sci.hokudai.ac.jp

2009/08/28 Sapporo

1 はじめに

近年の統計学の発展はめざましい．その原動力は計算機の性能の大幅かつ急速な進歩である．それによって統計学の方法は一新され，適用範囲も広がった．一方，大量の計算が必要なため，かっては実用的でないとされていた方法が復活し，大幅に適用範囲を広げたたこともある (Fisher の並べ替え検定など)．とくに 1990 年代に「MCMC 革命」とでもいうべき大変革が起こった．ここで MCMC とは Markov chain Monte Carlo 法の略である．原型は 1950 年代にすでに物理の分野 (多重積分の数値計算) にあった．もともと汎用性のある方法だったので，統計と結び付くのは自然なことであった．その余波としてベイズ統計学の復興があった．ベイズ統計など，30 年ほど前には，日本ではあまり研究されていなかったし評価も低かった．[4]

統計学への応用として，次のような代数の分野が目に付く．

(1) MCMC 法による分割表の大量生成方式へのグレブナー基底の応用．関連して多項式環や代数幾何．Strumfels [6] [1] 参照．

(2) 有限群上のランダムウォークへの表現論の応用．これについては Diaconis [7] と Ceccherini-Silberstein [8] がよい．

実は (1) と (2) は密接に関係している．(1) の問題点は，収束性の理論がないと思われることである．また (2) は統計学というより確率過程 (ランダムウォーク) の分野に属する．ところで，分割表の集合は対称群の Young 部分群による両側剰余類の集合と一対一に対応している．したがって，MCMC 法で分割表の大量生成しようとするなら，対称群の元の大量生成をすればよい．互換の集合は対称群の生成元になっているので，互換をランダムにとって次々に掛けていけば，対称群の元，したがって分割表がいくらでも得られる．しかし不思議なことに，(1) と (2) の関係はこれまで知られていないようである．

代数を使った統計の分野は「代数統計」とか「計算代数統計」と呼ばれている．代数の分野では，群論 (有限群，線形群，表現)，代数幾何 (多項式環)，可換環論 (対称式，不変式)，組合せ論 (Young 図形，数え上げ，母関数，有限幾何，グラフ，結合的概型) といったものが使えそうである．今この分野は第二期の爆発的発展の直前にあるように感じる (第一期はグレブナー基底の登場)．

2 データセットと分割表

データセットとは，有限集合の間の写像 $[h : N \to \mathcal{K}]$ のことである．単なる写像と区別するために括弧でくくってある．$n = |N|$ を**サイズ**という．統計学的には，N はサンプルの集合，$\mathcal{K}$ は名義尺度，h は確率変数で，$h(i)$ は $i \in N$ の属するカテゴリー (統計的な意味) である．すなわち，$\mathcal{K}$ 上のデータセッ

トは $\mathcal{K}$ 上の有限集合，すなわちスライスカテゴリー $\mathbf{set}/\mathcal{K}$ の対象である (**set** は有限集合と写像のなすカテゴリー)．したがって $\mathcal{K}$ 上のデータセット同士の**射** $\theta : [h : N \to \mathcal{K}] \to [h' : N' \to \mathcal{K}]$ は，写像 $\theta : N \to N'$ で，$h' \circ \theta = h$ を満たすものである．カテゴリー論の術語を使ってデータセットについて同型射，全射，単射，等化などの概念が定義できる．

スライスカテゴリーにおける直既約な対象は，$i_\kappa : \{*\} \to \mathcal{K}; * \mapsto \kappa$ (ここで $\kappa \in \mathcal{K}$) の形をしている．したがって κ に対応する Burnside 準同型は

$$\mathsf{tab}_\kappa : [h : N \to \mathcal{K}] \mapsto |\operatorname{hom}([i_\kappa],[h])| = |h^{-1}(\kappa)|$$

で与えられる．したがって Burnside 準同型は写像

$$\mathsf{tab} : \mathbf{set}/\mathcal{K} \to \mathbb{N}_0^{\mathcal{K}}; [h] \mapsto \mathsf{tab}_\kappa([h]) = (|h^{-1}(\kappa)|)_\kappa$$

を与える．配列 $\mathsf{tab}[h] = (|h^{-1}(\kappa)|)_\kappa$ をデータセット $[h]$ の**度数分布表** (table) という．こう考えると抽象バーンサイド環の理論が使える．

$$\mathsf{TAB}(n;\mathcal{K}) := \{\boldsymbol{c} = (c_\kappa)_{\kappa\in\mathcal{K}} \mid c_\kappa \in \mathbb{N}_0, \textstyle\sum c_\kappa = n\}$$

と置くと，上の写像 tab は，$\mathsf{DS}(N;\mathcal{K}) \to \mathsf{TAB}(n;\mathcal{K})$ (ここで $n := |N|$) を与える．

$\boldsymbol{c} \in \mathsf{TAB}(n;\mathcal{K})$ に対し，

$$\mathsf{DS}(\boldsymbol{c}) := \mathsf{tab}^{-1}(\boldsymbol{c}) = \{[h : N \to \mathcal{K}] \mid \mathsf{tab}[h] = \boldsymbol{c}\}$$

と置く．N 上の対称群 S_N は $\mathsf{DS}(N;\mathcal{K})$ に，右から合成によって作用する：$[h : N \to \mathcal{K}]\sigma := [h \circ \sigma]$.

Lemma 2.1 N はサイズが n とする．
(1) 2つのデータセット $[h : N \to \mathcal{K}]$, $[h' : N' \to \mathcal{K}]$ について次の条件は同値である：
(a) スライスカテゴリー $\mathbf{set}/\mathcal{K}$ において $[h] \cong [h']$.
(b) 全単射 $\pi : N \to N'$ があって，$h = h' \circ \pi$.
(c) $\mathsf{tab}[h] = \mathsf{tab}[h']$.
(2) $\mathsf{tab} : \mathsf{DS}(N;\mathcal{K}) \to \mathsf{TAB}(n;\mathcal{K}); [h] \mapsto \mathsf{tab}[h]$ は，全単射

$$\mathsf{DS}(\boldsymbol{c})/S_N \cong \mathsf{TAB}(N;\mathcal{K})$$

を誘導する．とくに $\mathsf{DS}(\boldsymbol{c})$ は S_N-軌道になっている．
(3) $\boldsymbol{c} = (c_\kappa) \in \mathsf{TAB}(N;\mathcal{K})$ で $[h] \in \mathsf{DS}(\boldsymbol{c})$ とする．このとき

$$S_h \backslash S_N \to \mathsf{DS}(\boldsymbol{c}); S_h\pi \mapsto [h \circ \pi] \tag{2.1}$$

は全単射である．ここで S_h は $[h] \in \mathbf{set}\mathcal{K}$ の自己同形群である．さらに

$$S_h := \{\pi \in S_N \mid h \circ \pi = h\} \tag{2.2}$$

(4) $|\mathsf{DS}(\boldsymbol{c})| = n! \Big/ \prod_{\kappa\in\mathcal{K}} c_\kappa!$

注意. (1) S_h は Young 部分群である．対応する $n = |N|$ の分割は $\mathsf{tab}[h]$.
(2) この程度のことにカテゴリーを持ち出すこともないのだが，$\mathcal{K}$ が順序集合のように構造を持っている場合，上のような一般論を用意しておくのが便利である．

2次元データセットは，(1次元) データセットの対 $([f : N \to \mathcal{L}], [g : N \to \mathcal{M}])$ (普通は $[f,g]$ のように書く) のことである．$\mathsf{DS}(N;\mathcal{L},\mathcal{M})$ によってそのような2次元データセットの集合を表す．$[f,g]$ と $[\langle f,g\rangle : N \to \mathcal{L}\times\mathcal{M}]$ を対応させることにより，2次元データセットは1次元データセットと1対1に対応する：

$$\mathsf{DS}(N;\mathcal{L},\mathcal{M}) \cong \mathsf{DS}(N;\mathcal{L}\times\mathcal{M}]$$

対称群の直積 $S_N \times S_N$ が $\mathsf{DS}(N;\mathcal{L},\mathcal{M}]$ に作用する：$[f,g](\sigma,\tau) := [f\circ\sigma, g\circ\tau]$．他方1次元データセットへの S_N の作用を考えると，$[f,g]\pi := [f\circ\pi, g\circ\pi]$ が y となっているが，これは対角作用である: $[f,g]\pi = [f,g](\pi,\pi)$．しかも $[f,g],[f',g'] \in \mathsf{DS}(N;\mathcal{L},\mathcal{M})$ が同型であることと，それらが S_N の対角作用により同じ軌道に含まれることは同値である．

$\mathsf{TAB}(n;\mathcal{L},\mathcal{M}) := \mathsf{TAB}(n;\mathcal{L}\times\mathcal{M})$ と書く．$\boldsymbol{x} = (x_{\lambda,\mu}) \in \mathsf{TAB}(n;\mathcal{L},\mathcal{M})$ に対し，

$$x_{\lambda+} := \sum_{\mu\in\mathcal{M}} x_{\lambda\mu},\ x_{+\mu} := \sum_{\lambda\in\mathcal{L}} x_{\lambda\mu},$$
$$x_{++} := \sum_{\lambda,\mu} x_{\lambda\mu} = \sum_{\lambda} x_{\lambda+} = \sum_{\mu} x_{+\mu}$$

と置く．$n := x_{++}$ を $\boldsymbol{x}$ の**サイズ**という．$(x_{\lambda+}) \in \mathsf{TAB}(n;\mathcal{L})$ と $(x_{+,\mu}) \in \mathsf{TAB}(n;\mathcal{M})$ を**周辺分布** (marginal distribution) という．次の可換図式を得る：

$$\begin{array}{ccc} & & \mathsf{DS}(N;\mathcal{L},\mathcal{M}) \\ & & \| \\ \mathsf{DS}(N;\mathcal{L}\times\mathcal{M}) & \xrightarrow{\cong} & \mathsf{DS}(N;\mathcal{L})\times\mathsf{DS}(N;\mathcal{M}) \\ \downarrow{\scriptstyle \mathsf{tab}_{\mathcal{L}\times\mathcal{M}}} & & \downarrow{\scriptstyle \mathsf{tab}_{\mathcal{L}}\times\mathsf{tab}_{\mathcal{M}}} \\ \mathsf{TAB}(n;\mathcal{L}\times\mathcal{M}) & \xrightarrow{\mathsf{mar}} & \mathsf{TAB}(n;\mathcal{L})\times\mathsf{TAB}(n;\mathcal{M}) \end{array}$$

ここで $\mathsf{mar} : (x_{\lambda\mu}) \mapsto ((x_{\lambda+}),(x_{+\mu}))$ である．また $\mathsf{tab}[f,g] = (|f^{-1}(\lambda)\cap g^{-1}(\mu)|)_{\lambda,\mu}$ はデータセット $[f,g]$ の**分割表** (contingency table) という．

$(\boldsymbol{a},\boldsymbol{b}) \in \mathsf{TAB}(n;\mathcal{L})\times\mathsf{TAB}(n;\mathcal{M})$ に対し

$$\mathsf{DS}(\boldsymbol{a},\boldsymbol{b}) := \mathsf{mar}^{-1}(\boldsymbol{a},\boldsymbol{b})$$

と置く．

Lemma 2.2 $[f,g],[f',g'] \in \mathsf{DS}(N;\mathcal{L}\times\mathcal{M})$ で $|N| = n$ とする．
(1) $[f,g],[f',g']$ が同じ周辺分布を持つ，すなわち $\mathsf{mar}[f,g] = \mathsf{mar}[f',g']$)，ための必要十分条件は，$\sigma,\tau \in S_N$ で，$f = f'\sigma$, $g = g'\tau$ を満たすものが存在することである．とくに $S_N\times S_N$ は $\mathsf{DS}(,\boldsymbol{b})$ に可移に作用する．

(2) $\mathsf{tab}[f,g] = \mathsf{tab}[f',g']$ であるための必要十分条件は，$\pi \in S_N$ で $f = f'\pi, g = g'\pi$ を満たすものが存在すること *(*同じことだが，**set**$/\mathcal{L}\times\mathcal{M}$ において同型なこと*)* である．
(3) $\sigma,\tau,\pi \in S_N$ に対し，

$$\begin{aligned} \mathsf{tab}[f\circ\pi, g\circ\pi] &= \mathsf{tab}[f,g], \\ \mathsf{tab}[f\circ\sigma, g\circ\tau] &= \mathsf{tab}[f\circ\sigma\tau^{-1}, g] \end{aligned}$$

Corollary 2.3 $[f,g] \in \mathsf{DS}(N;\mathcal{L},\mathcal{M})$ に対し $\pi \mapsto [f\pi,g]$ と $(\sigma,\tau)\mapsto[f\sigma,g\tau]$ は次の一対一対応を引き起こす：

$$S_{f_0}\backslash S_N/S_{g_0} \xrightarrow{\cong} (S_f\times S_g)\backslash(S_N\times S_N)/S_N^{\mathrm{diag}} \xrightarrow{\cong} \mathsf{TAB}(\boldsymbol{a},\boldsymbol{b})$$

Corollary 2.4 (1) $\mathsf{tab}\times\mathsf{tab}$ による $\mathsf{DS}(\boldsymbol{a},\boldsymbol{b})$ 上の一様分布の像は超幾何分布である：

$$\mathrm{Prob}(\mathsf{tab}[f,g] = \boldsymbol{x}) = \frac{\boldsymbol{a}!\,\boldsymbol{b}!}{n!\,\boldsymbol{x}!} =: H(\boldsymbol{x})$$

ここで $\boldsymbol{x}! = \prod_{\lambda\mu} x_{\lambda,\mu}!$ など．
(2) $[f,g] \in \mathsf{DS}(\boldsymbol{a},\boldsymbol{b})$ とする．このとき $(\sigma,\tau)\mapsto \mathsf{tab}[f\sigma,g\tau]$ による $S_N\times S_N$ 上の一様分布の像は同じ超幾何分布である．同じことだが，

$$\frac{1}{n!^2}\#\{(\sigma,\tau) \mid \mathsf{tab}[f\sigma,g\tau] = \boldsymbol{x}\} = H(\boldsymbol{x})$$

(3) $\sigma \mapsto \mathsf{tab}[f\sigma,g]$ による S_N 上の一様分布の像も同じ超幾何分布である．

$N = \{1,2,\cdots,n\}$, $\boldsymbol{a} \in \mathsf{TAB}(n;\mathcal{L}), \boldsymbol{b} \in \mathsf{TAB}(n;\mathcal{M})$ とする．$[f,g]$ を $DS(\boldsymbol{a},\boldsymbol{b})$ から取るっておく．例えばこれは観測データから得られたデータセットでよい．対称群 S_n の元 $\sigma_1,\sigma_2,\cdots$ をランダムに取って行くと $\mathsf{tab}[f\sigma_1,g],\mathsf{tab}[f\sigma_2,g],\cdots$ は $TAB(\boldsymbol{a},\boldsymbol{b})$ のランダムウォークである．生起確率は超幾何分布にしたがう．これで分割表のサンプリングが得られる．対称群の元のランダムサンプリングが問題になるが，Diaconis の方法では，ランダムに互換 $\tau_1,\tau_2,\cdots$ を取ってそれを順に掛けて行く：$1,\tau_1,\tau_2\tau_2,\cdots$．ただし奇置換と偶置換が交互に入れ替わるので，このランダムウォークはいわゆるエルゴード性を満たしておらず，定常分布を持たない．幸いなことに $\sigma\mapsto\mathsf{tab}[f\sigma,g]$ で誘導される分割表のウォークは，まれな例外 (周辺度数 $x_{\lambda+},x_{+\mu}$ がつねに 0 か 1) を除いてエルゴード性を満たし，したがって超幾何分布に収束するランダムウォークである．

ランダムな互換 τ によって $[f,g]$ から $[f,\tau,g]$ を作ることは，分割表 $\mathsf{tab}[f,g]$ のレベルでいうと，$\mathsf{tab}[f,g]$ に $B = \boxed{\begin{smallmatrix} -1 & 1 \\ 1 & -1 \end{smallmatrix}}$ を加えること (非負成分は 4 つ) である．したがってこの遷移法則は完全に分割表だけで記述できる．もとのデータセットの取り方に依らない．

この行列 B は，MCMC 法の有名な行列 (マルコフ基底，基本移動 basic move) である．あとで見るように収束は遅い (収束率 $(n-3)/(n-1)$ は 1 に

近い)．不思議なことに，対称群上の RW から誘導される分割表たちの RW について書いた文献が見つからない．Diaconis など両方の研究で重要な成果を上げているのにである．しかし考えをさかのぼれば，両者を結びつける方法はフィッシャーの並べ替え検定 (permutation test) そのものである [10]．フィッシャーの時代 (1930 年代) は計算機もなく実用性はほとんどなかった．今では再評価されている．このようなリサンプリング法の流れの先にブートストラップ法がある．代数的に解釈するなら，並べ替え検定で対称群を対称半群に取り替えものである．

今から 30 年ほど前，比較言語学のオズワルドのシフト検定法のことを聞いて，置換群が使えることに気づいた．数年前に，並べ替え検定や，分割表列挙への MCMC 法の適用などを知るようになり，ほとんど公表していない昔の研究との関係を思い出した．

3 群作用を伴う有限集合上のランダムウォーク

確率統計ではまったく標準的でない記号を使う．$\mathbf{\Delta}$ を単位区間とする：$\mathbf{\Delta} := \{r \in \mathbb{R} \mid 0 \le r \le 1\}$．さらに，$X$ 上の確率分布の集合を ∇X とする：

$$\nabla X := \left\{\mu : X \to \mathbf{\Delta} \,\middle|\, \textstyle\sum_{x\in X} \mu(x) = 1\right\}$$

また X を頂点集合とする単体を ΔX とする：

$$\Delta X := \left\{\widehat{\mu} := \textstyle\sum_{x\in X} \mu(x)x \mid \mu \in \nabla X\right\}$$

人によってはこの集合を**マルコフ空間** (Markov space) とか**確率単体**と呼んでいる．

マルコフ空間の間の**マルコフ写像** (ときに**マルコフ作用素**) とはアファイン写像のことである．したがってマルコフ写像 $\boldsymbol{P} : \Delta X \to \Delta Y$ は

$$\boldsymbol{P}(r\widehat{\mu} + (1-r)\widehat{\nu}) = r\boldsymbol{P}(\widehat{\mu}) + (1-r)\boldsymbol{P}(\widehat{\nu}),\ r \in \mathbf{\Delta}$$

を満たす．アファイン性によりマルコフ写像 $\boldsymbol{P}$ は頂点の行き先で決まる．

$$\begin{aligned}\boldsymbol{P} = \widehat{P} \quad &: \quad x \mapsto \sum_{y\in Y} P(x,y)\,y,\\ &: \quad \widehat{\mu} \mapsto \widehat{\mu P}\end{aligned}$$

ここで，$\mu = (\mu(x))$ を行ベクトル，$P = (P(x,y))$ を $X \times Y$ 型の行列と見なした．$P(x,y) \in \mathbf{\Delta}$ であり，$\sum_{y\in Y} P(x,y) = 1$ である．$x \mapsto P(x,-)$ はアファイン写像 $\nabla X \to \nabla Y$ を与える．

X 上の**マルコフ連鎖** (Markov chain) とは，$(\Delta X, \boldsymbol{\mu}_0, \boldsymbol{P})$ のことである．ここで，$\boldsymbol{\mu}_0 \in \Delta X$ で $\boldsymbol{P}$ はマルコフ作用素である．対応する行列 $P = (O(x,y))$ を**遷移行列** (transision matrix) という．同じことだがマルコフ連鎖を $(\nabla X, \mu_0, P)$ と書いてもよい．この場合 P は $nablaX$ 上のアファイン写像になっている．

$P(x,y)$ は $x \in X$ が $y \in Y$ に移動する確率を表す．したがって行列のベキ P^m の成分 $P^m(x,y)$ は，m ステップで x が y に移動する確率を表す．したがって ΔX で考えると，マルコフ連鎖とは，離散線形力学系 $\boldsymbol{P} : \Delta X \to \Delta X$ のことである．マルコフ連鎖の用語と力学系の用語は対応するものが多い．したがってマルコフ連鎖があればマルコフ空間の元の列

$$\boldsymbol{\mu}_0 \to \boldsymbol{\mu}_1 := \boldsymbol{P}(\boldsymbol{\mu}_0) \to \cdots \to \boldsymbol{\mu}_m := \boldsymbol{P}^m(\boldsymbol{\mu}_0) \to \cdots$$

が得られる．同じことだが，行列の言葉では確率分布の列

$$\mu_0 \to \mu_1 := \mu_0 P \to \cdots \to \mu_m := \mu_0 P^m \to \cdots$$

となる．確率過程の流儀にしたがって，行列のベクトルへの作用は右からとする．

結局のところ，マルコフ連鎖とは，係数体を単位区間 $\mathbf{\Delta}$ にした線形代数と言える．マルコフ連鎖の理論における**基本的な問題**は，極限 $\lim_{m\to\infty} \boldsymbol{P}^m(\boldsymbol{\mu}_0)$，同じことだが極限分布 $\lim_{m\to\infty} \mu_0 P^m$ が存在するかである．

G を有限群，X を右 G-集合とする．群作用 $X \times G \to X$ を線形に拡大すれば ΔG の ΔX への線形な作用 $\Delta X \times \Delta G \to \Delta X; (\boldsymbol{\mu}, \boldsymbol{P}) \mapsto \boldsymbol{\mu P}$ が得られる．$p \in \nabla G$ に対し

$$\widehat{p} := \sum_{g\in G} p(g)g \in \Delta G$$

と置けば，$\widehat{p}$ は ΔX 上のマルコフ作用素を誘導する：

$$\Delta_X(p):\Delta X\to\Delta X;\widehat{\mu}=\sum\mu(x)x\mapsto\widehat{\mu}\widehat{p}=\widehat{\mu*p}$$

ここで

$$(\mu*p)(x)=\sum_x\mu(xg^{-1})p(g).$$

したがってマルコフ連鎖 $(\Delta X,\mu_0,\Delta_X(p))$ が得られる．このような群の作用する集合上のマルコフ連鎖を**ランダムウォーク** (以下 RW) と呼ぶ．この RW の遷移確率行列は

$$P_X(x,y):=\sum_{g:xg=y}p(g)$$

で与えられる．ここで，和は $xg=y$ を満たす $g\in G$ について取る．

$p,q\in\nabla G$ に対し，合成積を

$$(p*q)(g)=\sum_{h\in G}p(h)q(h^{-1}g)$$

で定義すれば，

$$\widehat{pq}=\widehat{p*q},\ \Delta_X(p)\circ\Delta_X(q)=\Delta_X(p*q)$$

である．初期分布 $\mu_0=\widehat{\mu}_0$ から始まる分布の列は，$\mu_m=\mu_0\widehat{p}^m$，あるいは $\mu^{(m)}=\mu_0P_X^m$ で与えられる．$\widehat{p}\in\Delta G$ に対応する行列は $P(g,h):=p(g^{-1}h)$ で与えられる．これを使うなら，$\mu^{(m)}=\mu_0P^m$ とも書ける．したがって μ_m の漸近的挙動は主に $\widehat{p}\in\Delta G\subset\mathbb{C}G$ に依存する．

群上の RW を考えるとき，これまでは ∇G の中で考えることが普通だったが，マルコフ空間で考える方がはるかに簡潔である．

例．$G=X=C_N=\langle g_0\rangle$ 奇数位数 N の巡回群．

$$p(g)=\begin{cases}1/2 & g=g_0^{\pm1}\\ 0 & \text{else,}\end{cases}\qquad P(g,h)=\begin{cases}1/2 & gh^{-1}=g_0^{\pm1}\\ 0 & \text{else}\end{cases}$$

$$p^{(*m)}(g_0^j)=\frac{1}{N}\sum_{k=0}^{N-1}\left(\cos\frac{2k\pi}{N}\right)^m\left(\cos\frac{2jk\pi}{N}\right),$$

$P(g_0^i,g_0^j)=p(g_0^{j-i})$.

$\therefore P^m\to u_G$ (Uniform distribution). Diaconis LN, 1988.

4 スペクトル分解

$p\in\nabla G$ が類関数の場合を考える．この場合

$$\widehat{p}=\sum_{g\in G}p(g)g\in(\Delta G)^G=\nabla G\cap Z(\mathbb{C}G)$$

である．したがって有限群の表現論が使える．$\mathrm{Irr}(G)$ を G の (複素) 既約指標の集合とする．表現論により

$$\widehat{p}=\sum_{\chi\in\mathrm{Irr}(G)}\omega_\chi(\widehat{p})\,e_\chi.$$

ここで

$$\omega_\chi:\mathbb{C}G\to\mathbb{C};g(\in G)\mapsto\frac{\chi(g)}{\chi(1)}$$

は線形写像で，したがって

$$\omega_\chi(\widehat{p})=\sum_{g\in G}p(g)\frac{\chi(g)}{\chi(1)}=\frac{|G|}{\chi(1)}\langle\chi,p\rangle$$

である．さらに e_χ は中心原始ベキ等元である：

$$e_\chi:=\frac{\chi(1)}{|G|}\sum_{g\in G}\overline{\chi(g)}\,g\in Z(\mathbb{C}G)$$

このことから

$$\widehat{p}^m=\widehat{p^{(*m)}}=\sum_{\chi\in\mathrm{Irr}(G)}\omega_\chi(\widehat{p})^m\,e_\chi$$

Theorem 4.1 $S=supp(p):=\{g\in G\mid p(g)\neq0\}$ とし，N を $S^{-1}S$ で生成された *(正規)* 部分群とする．

(1) $m\to\infty$ で $p^{(*m)}\to u_G$ *(一様分布)* となるための必要十分条件は，$N=G$ となることである．

(2) $g_0\in S$ とし，$q(g):=1/|N|(q\in g_0N);:=0(else)$ とする．このとき $m\to\infty$ で，

$$p^{(*m)}-q^{(*m)}\approx C\rho^m,\quad \rho:=\max_{\chi,g}\left|\frac{\chi(g)}{\chi(1)}\right|$$

ここで C は定数，χ は $|\chi(g)|<\chi(1)$ を満たす既約指標を動く．

証明は指標理論から容易. $|\omega_\chi(\widehat{p})| \le 1$ で，等号成立が $N \subset \mathrm{Ker}(\chi)$ のときに限ることを使う. (1) は Diaconis も証明なしでどこかに書いている. 結局類関数 p から得られる有限群上の RW は，一様部分と収束部分，振動部分に分かれることが分かった.

例. $G = S_n$, $\tau \in C$ 互換のなす共役類とする..

$$p(g) = \begin{cases} \frac{2}{n(n-1)} & g \in C \\ 0 & \text{else} \end{cases}$$

この場合，一様部分 $\mathrm{U} = (1/n!)1_G$, 振動部分 $\mathrm{O} = (-1)^n(1/n!)\mathrm{sgn}$ となっている. $\chi_\lambda \neq 1_G, \mathrm{sgn}$ のとき,

$$\rho_\lambda = \frac{\chi_\lambda(\tau)}{\chi_\lambda(1)} = \frac{1}{n(n-1)}\sum_i i^2(\mu_i - \mu_i')$$

ただし $\lambda = 1^{\mu_1}2^{\mu_2}\cdots$ で，λ' は分割 λ の共役である. $|\rho_\lambda|$ が 1 に近いため，収束は遅い.

$\rho := \max(|\rho_\lambda| : \chi_\lambda(1) > 1) = \frac{n-3}{n-1} < 1$ である. 等号成立は $\lambda = [n-1,1], [2,1^{n-2}]$ のとき.

$$p^{(m)} \approx \frac{1}{n!}(1_G + (-1)^m \mathrm{sgn})$$
$$\therefore \quad P^{(2m)} \to \mathrm{U}_{A_n},\ P^{2m+1} \to \mathrm{U}_{S_n - A_n}$$

C が n-サイクルからなるなら $\rho = 1/(n-1)$ なので，収束はきわめて速い. ただし実用的ではなさそう.

問題 有限群 G に対し，決定せよ:
$\rho = \mathrm{Min}_t\mathrm{Max}_\chi(|\chi(t)/\chi(1)| : |\chi(t)| < \chi(1))$

対称群 S_n の場合, $\rho = 2/n(n-3)$ が $(n-1)$-サイクルと $[n-2,2], [2^2, 1^{n-4}]$ によって実現. .

5 ヘッケ環

K を有限群 G の部分群, $p \in (\nabla G)^K$ (K-類関数) とする. $\widehat{p} := \sum_{g\in G} p(g)g \in \Delta G \cap (\mathbb{C}G)^K$ である. さらに X を右有限 G-集合とする. このとき

$$(\Delta X)^K \times (\Delta G)^K \to (\Delta X)^K; (\widehat{\mu}, \widehat{p}) \mapsto \widehat{\mu * p} = \widehat{\mu}\widehat{p}$$

$(\Delta X)^K \cong \Delta(X/K)$ なので, $\widehat{p}$ は任意の μ_0 から始まる X/K 上の RW を誘導する. 遷移確率は

$$\nabla_{X/K}(xK, yK) = p(x\backslash yK) := \sum_{xg \in yK} p(g)$$

である. $e_K := (1/|K|)\sum_{k\in K} k$ を $K \le G$ に対応する $\mathbb{C}G$ のベキ等元とする. $\mathcal{H}(G,K) := e_K\mathbb{C}Ge_K \subset \mathbb{C}G$ を**ヘッケ環**という単位元は e_K である. 対応 $(\Delta X)^K \cong \Delta(X/K)$ は $ge_K \leftrightarrow gK$ から誘導されている.

$$(\Delta G)^K \xrightarrow{\eta} \Delta\mathcal{H}(G,K) \to \Delta(X/K, X/K)$$

を考える. ここで $\Delta G)^K$ はマルコフ空間の K-不変元全体のなす Δ-多元環, $\Delta\mathcal{H}(G,K) := e_K\mathbb{C}Ge_K \subset \mathcal{H}(G,K)$. このとき

$$\eta : \widehat{p} \mapsto e_K\widehat{p}e_K = \widehat{p}e_K = e_K\widehat{p}$$

は全射である.

$\widehat{p}$ から誘導押される X/K 上の RW を考えるのだが，中心化環 $(\mathbb{C}G)^K$ と $\mathcal{H}(G,K)$ は半単純なので，状況は群上の RW と良く似ている. 実際，もし $\eta(\widehat{p}) = e_K\widehat{p} \in Z(\mathcal{H}(G,K)$ ならば,

$$\widehat{p} = \sum_{\chi \le 1_K^G} \omega_{K,\chi}(\widehat{p})e_Ke_\chi \tag{5.3}$$

と分解できる. ここで $\omega_{K,\chi} : \mathcal{H} := \mathcal{H}(G,K) \to \mathbb{C}$ は線形写像で

$$\omega_{K,\chi} : e_Kge_K \mapsto \chi(e_Kge_K)/\chi(e_K)$$

で定義されるものである.

$$\omega_{K,\chi}(\widehat{p}) = \sum_{g\in G} p(g)\frac{\chi(e_Kge_K)}{\chi(e_K)}$$

である. このとき $\omega_{K,\chi}$ は $Z(\mathcal{H}) \to \mathbb{C}$ なる多元環準同型である. より一般に $a \in \mathcal{H}$ と $b \in Z(\mathcal{H})$ に対し $\omega_{K,\chi}(ab) = \omega_{K,\chi}(a)\omega_{K,\chi}(b)$ である.

(5.3) より

Theorem 5.1

$$\widehat{p}^m = \widehat{p^{(*m)}} = \sum_{\chi \le 1_K^G} \omega_{K,\chi}(\widehat{p})^m e_Ke_\chi \tag{5.4}$$

とくに $S := suppp$, $N := \langle S^{-1}S\rangle$ とする. このとき N が X/K に可移に作用すれば,

$$\nabla_{X/K}(p)^{(*m)} \to u_{X/K} \quad (\text{一様分布})$$

分割表の場合は，ゲルファント対 $(G\times G, G^{\mathrm{diag}})$ が現れた．その場合，G は対称群であった．つまり $\mathcal{H}(H\times G, G^{\mathrm{diag}})$ は可換である．X, Y を有限右 G-集合とすれば，$X\times_G Y := X\times Y/\sim$ (ここで $(xg, yg)\sim(x,y)$ $(\forall g\in G)$．マルコフ空間のテンソル積を

$$\Delta X\otimes_G \Delta(Y) = \Delta(X\times Y)^{G^{\mathrm{diag}}} = \Delta(X\times Y/\sim)$$

と定義する．

$r\in\nabla(X\times Y)^{G^{\mathrm{diag}}}$ とする．すなわち $r(g^u, h^u) = r(g,h)$ $(\forall g,h,u\in G)$．このとき $(X\times_G Y$ 上のマルコフ作用素

$$\widehat{r} : x\otimes y\mapsto \sum_{g\in G} r'(g)xg\otimes y \quad \text{ここで}$$
$$r'(g) := \sum_{h\in G} r(gh,h), \quad r'\in(\nabla G)^G$$

が得られる．

$\nabla(G\times G)^{G^{\mathrm{diag}}}\cong(\nabla G)^G$ に注意すると

$$(x\otimes y)\widehat{r'} = x\widehat{r'}\otimes y = (x\otimes y)\widehat{r}$$

したがって $\widehat{r} = \widehat{r'}$ となる．結局分割法のときと同じ状況になっている．よく知られているように，この場合はヘッケ環は普通の $\mathbb{C}G$ の表現と同じである．

$X = \mathsf{DS}(\boldsymbol{a}), Y = \mathsf{DS}(\boldsymbol{b})$ のときが分割表の場合であり，実際 $X\times_G Y\cong\mathsf{TAB}(\boldsymbol{a},\boldsymbol{b})$ となっている．

一般のゲルファント対の場合にこれまでの議論を拡張することは可能である．しかしすでに締め切りも大幅にすぎたので，一般的なゲルファント対の場合と，系統樹のサンプリングへの応用，それに分割法の収束の議論は別の機会に述べたい．

一般的な文献については，吉田知行の神戸大学での講演報告にある．この分野でもっとも活発な活動をしているのは，Diaconis のグループであろう．日本でも竹村，青木などの研究がある．

最後に講演の機会を与えていただいたオーガナイザー，とくに森田英章氏に感謝します．

参考文献

[1] 日比孝之『グレブナー基底の現在』数学書房 2006

[2] 伊庭，種村『計算統計 2— マルコフ連鎖モンテカルロ法とその周辺』岩波 2005

[3] 伊庭幸人『ベイズ統計と統計物理』岩波書店 2003

[4] 中山伊知郎 (編)『現代統計学大辞典』東洋経済新報社 1980

[5] Diaconis, P. and Gangolli, A. , Rectangular arrays with fixed margins, in "Discrete Probability and Algorithms (D. Aldous et al., eds.)", 15–41, Springer, New York ,1995.

[6] L.Prachter, B.Sturmfels (編), "Algebraic Statistics for Computational Biology," Cambridge, 2005

[7] P.Diaconis, "Group Representaions in Probability and Statistics," LN-Monograph series 11, Institute of Math.Stat., 1988.

[8] T. Ceccherini-Silberstein, ect.,"Harmonic Analysis on Finite Groups," Cambridge, 2008.

[9] L.Salloff-Coste, Random Walks on Finite Groups, in "Probability on Discrete Structures" (Encyclopaedia of Mathematical Sciences), 264–346, 2003.

[10] P.W.Mielke, K.J.Berry, "Permutation Methods", Springer, 2001, 2007.

[11] 渡部隆一『マルコフ点チェーン』数学ワンポイント双書 31，共立出版 1979.

[12] 吉田知行「分割表の一致率検定への有限群論と組合せ論の応用」代数学シンポジウム (神戸大学 2007) 報告集.